

This listing of claims will replace all prior versions and listings of claims in the application.

Listing of Claims:

1. (Currently amended) A document digitizing method of digitizing a document in printed form, comprising:
 - optically scanning the document;
 - forming and storing a digitized image file from the optically scanned document;
 - optically recognizing under computer control characters in the optically scanned document; and
 - forming and storing a text file of the optically recognized characters in the document, wherein each digitized image file is correlated with a corresponding text file and wherein the digitized image file and text file have common file names and are distinguished by appropriate file extensions.
2. (Original) The method of claim 1 in which the document includes plural pages and a separate digitized image file is formed for each page of the document.
3. (Original) The method of claim 2 in which a separate text file is formed for each page of the document.
4. (Original) The method of claim 1 in which the document includes plural pages and a separate text file is formed for each page of the document.
5. (Cancelled)
6. (Cancelled)
7. (Currently amended) The method of claim ~~5~~1 in which corresponding digitized image files and text files are correlated by a mapping table or algorithm.

8. (Original) The method of claim 1 further comprising retrieving a digitized image file for a document based upon a text string in the text file corresponding to the digitized image file.
9. (Original) The method of claim 1 in which the digitized image file is compressed and of a lossless image file format.
10. (Original) The method of claim 1 in which the text file is of a simplified file format based upon ASCII characters.
11. (Original) The method of claim 1 in which optical character recognition is applied to all text characters in the optically scanned document.
12. (Currently amended) In a document digitizing system for optically scanning and digitizing a document in printed form and having a computer readable medium, a digitized document data structure for digitized documents stored in the computer readable medium, comprising:
 - a digitized image file representing a digitized image from the optically scanned document; and
 - a text file of text characters optically recognized from the optically scanned document, wherein each digitized image file is correlated with a corresponding text file and wherein the digitized image file and text file have common file names and are distinguished by appropriate file extensions.
13. (Cancelled)
14. (Cancelled)
15. (Original) The data structure of claim 12 in which the document includes plural pages, the data structure comprising a separate digitized image file for each page of the document and a corresponding separate text file for each page of the document.

16. (Original) The data structure of claim 12 in which the digitized image file is compressed and of a lossless image file format.
17. (Original) The data structure of claim 12 in which the text file is of a simplified file format based upon ASCII characters.
18. (Currently amended) In a document digitizing system for optically scanning and forming a digitized image file of a document having text characters in printed form, a method of retrieving the digitized image file for a document, comprising:
storing digitized image files for plural printed documents in association with text files of the text characters in each document, the text files being generated by computer optical character recognition of the digitized image files or related image files, wherein the associated digitized image files and text files have common file names and are distinguished by appropriate file extensions;
searching the text files to identify any having a selected text string; and
providing access to the digitized image files corresponding to the text files identified as having the selected text string.
19. (Original) The method of claim 18 in which providing access to the digitized image files includes allowing a user to selectively display any of the digitized image files corresponding to the text files identified as having the selected text string.
20. (Cancelled)
21. (Original) The method of claim 18 in which searching the text files to identify any having a selected text string includes specifying multiple separate text strings and searching the text files in a batch to identify any text files having any of the separate text strings.
22. (Original) The method of claim 18 in which the text files have file names, the method further comprising storing the file names of the text files identified as having the selected text string.

23. (Presently amended) ~~In~~ A computer-readable medium encoded with computer executable instructions to perform a document digitizing ~~software~~ method for digitizing a document in printed form, comprising:

- ~~software for~~ optically scanning the document;
- ~~software for~~ forming and storing a digitized image file from the optically scanned document;
- ~~software for~~ optically recognizing under computer control characters in the optically scanned document; and
- ~~software for~~ forming and storing a text file of the optically recognized characters in the document, wherein each digitized image file is correlated with a corresponding text file and wherein the digitized image file and text file have common file names and are distinguished by appropriate file extensions.

24. (Cancelled)

25. (Cancelled)

26. (Presently amended) The computer-readable medium of claim 23 further comprising ~~software~~ executable instructions for retrieving a digitized image file for a document based upon a text string in the text file corresponding to the digitized image file.

27. (New) In a document digitizing system for optically scanning and forming a digitized image file of a document having text characters in printed form, a method of retrieving the digitized image file for a document, comprising:

- storing digitized image files for plural printed documents in association with text files of the text characters in each document, the text files being generated by computer optical character recognition of the digitized image files or related image files;
- searching the text files to identify any having a first text string;
- searching the text files that have the first text string to identify any having a second text string; and

DOCKET NO.: MSFT-3979/160226.1
Application No.: 09/862,728
Office Action Dated: June 18, 2004

PATENT

providing access to the digitized image files corresponding to the text files identified as having the first and second text strings.